Wireline Competition Bureau Staff Report

Possible Methodologies for Establishing Reasonably Comparable Broadband Rates for Fixed Services

June 30, 2014

Introduction

In the *USF/ICC Transformation Order*, the Commission required that as a condition of receiving Connect America Fund support, recipients must offer voice and broadband services in supported areas at rates that are reasonably comparable to rates for similar services in urban areas.[1] The Commission concluded that rural rates for broadband service would be deemed "reasonably comparable" to urban rates if those rates "fall within a reasonable range of the national average urban rate for broadband service."[2] It delegated authority to the Wireline Competition and Wireless Telecommunications Bureaus to conduct an annual survey of urban broadband rates in order to derive a national range of rates for broadband service.[3] In the *USF/ICC Transformation FNPRM*, the Commission sought comment on whether using two standard deviations would be the appropriate methodology for determining reasonable comparability, or should another methodology be used.[4]

The Wireline Competition Bureau (Bureau) is working to develop an approach for determining an upper range of rates that could be reasonably comparable to urban broadband prices for a broadband service with characteristics similar to a specified minimum download speed, upload speed and usage allowance. Our objective is to develop an approach that is flexible enough to take account any changes the Commission may make in the future regarding broadband performance obligations for recipients of Connect America funding.[5]

Developing a methodology for setting a reasonably comparable broadband benchmark involves (1) defining terms and scope based on the *USF/ICC Transformation Order*, (2) creating a sampling plan, (3) processing the collected data, and (4) analyzing the data. We explain below each step in this process, specifying the decisions that the Bureaus have already made regarding the execution of the urban rate survey and identifying the options for analyzing the data that has been collected.

---

[1] *Connect America Fund et al.*, WC Docket No. 10-90 et al., Report and Order et al., 26 FCC 17663, 17693, 17695, paras. 81, 86 (2011); *aff'd sub nom. In re: FCC 11-161*, 2014 WL 2142106 (10th Cir. May 23, 2014) (*USF/ICC Transformation Order*). *See also* 47 U.S.C. § 254(b). Recipients are also required to submit pricing data for both their voice and broadband offerings with their annual reports. 47 C.F.R. § 54.313(a)(7); *USF/ICC Transformation Order*, 26 FCC Rcd at 17856, para. 594.

[2] *USF/ICC Transformation Order*, 26 FCC Rcd at 17708, para. 113.

[3] *Id*. at 17708-09, para. 114.

[4] *Id*. at 18046-47, para. 1026. The Commission also asked whether it would be appropriate to use prices normalized to disposable income as a method for determining reasonable comparability, and whether rates should be presumed reasonably comparable if a given provider is offering the same rates, terms and conditions (including capacity limits, if any) to both rural and urban customers. *Id*.

[5] *Connect America Fund*, WC Docket No. 10-90, Order, 28 FCC Rcd 4242, 4248-49, para. 22 (Wireline Comp. Bur. 2013) (*Urban Rate Survey Order*) (noting that the Commission intends to monitor trends in marketplace and expects Connect America Fund obligations to evolve).

Implementation of the Survey

Definitions

In 2013, the Bureaus adopted the form and content of the urban rate survey. We decided to compute the "national average urban rate for broadband service" based on the mean of residential, non-promotional, advertised rates offered to potential new customers by firms in urban areas, i.e. list prices.[6] Given this, we designed a survey and methodology to estimate this parameter. The specific statistical interpretation used for development of the survey and estimation from the data collected is given in the Appendix.

The Bureaus made the decision not to create a national average urban rate that blends rates derived from fixed and mobile data.[7] Satellite broadband also was excluded from the sampling frame.[8] The Bureaus made the decision not to include existing contracts, but instead to collect rates only for new offered service. The Bureaus made the decision to collect rates on all standalone service plans offered to residential customers. As a result, in our sample, for each plan offered, the provider reported the advertised download bandwidth, the advertised upload bandwidth, the usage allowance (if any), and the monthly rate.

The Bureaus made a decision to define urban rates based on whether the rate was offered in an urban census tract.[9] A census tract was defined as urban if it contained any census-defined Urban Areas or Urban Clusters. Census tracts served as the geographic unit for which providers were asked to report residential broadband rates.

Survey Sample Selection

A sample of 500 survey units was randomly selected with replacement. These survey units were chosen by the Bureau's Industry Analysis and Technology Division (IATD) in a two-step process. First, 500 census tracts were randomly selected from all urban census tracts (as defined above). Second, for each of these selected census tracts a provider was chosen, using FCC Form 477 data.[10] This census tract-provider pair constitutes a sampling unit for which a survey was sent. Each of these sampling steps is explained below.

The frame for the selection of urban census tracts was provided by the Excel file "urbantracts_list _all.xls" which listed 58,331 urban census tracts encompassing the 50 states, the District of Columbia, and Puerto Rico. The first phase in the sample selection process was to randomly select, using household

[6] *Id.* at 4248, para. 20.

[7] *Id*. at 4243-44, at para. 6.

[8] *Id.*

[9] *Id*. at 4244-45, paras. 8-11.

[10] The Bureaus decided in the *Urban Rate Survey Order* to use FCC Form 477 data to select the urban providers of fixed voice and broadband services. *See Urban Rate Survey Order*, 28 FCC Rcd at 4244, para. 8. The Commission specifically directed the Bureaus to use Form 477 information to determine the sample for the voice rate survey, *see USF/ICC Transformation Order*, 26 FCC Rcd at 17694, para. 85, and it also noted that Form 477 contained necessary information on broadband providers, *see id*. at 17701, para. 103 n.168, 17755-56, 296 n.394.

weights, 500 census tracts with replacement[11] from this list of urban census tracts.[12] The selection was weighted proportionately by the number of households in the census tracts which was also provided in the file. The selection was performed using the "RandomChoice" function in Mathematica. The selection process produced an Excel file "urban tracts sample broadband.xls" of 498 unique census tracts; two census tracts were each selected twice.[13]

An Excel file ("broadband_v2") listing Fixed Broadband service providers reporting subscribers in the 498 unique census tracts in the sample was prepared based on Form 477 December 2012 filings. The file also gave the number of residential connections each provider had in each census tract in the sample.

For each of the 500 census tracts in the sample, a service provider was randomly selected from the providers of Fixed Broadband service for that census tract as listed in "broadband_v2" using the "RandomChoice" function in Mathematica. Because different providers in the same census tract may offer service to substantially different numbers of households, the selection was weighted based on the number of residential subscribers for each provider in the census tract as now described.

A service provider was given weight = 1 if the provider had more than 7% of the total residential subscribers in the census tract. Otherwise, the provider was given the weight = 1/(N+1) where N is the number of providers with 7% or less of the total residential subscribers in the census tract. So, if the census tract had only one service provider with 7% or less of the total residential subscribers in the census tract,[14] that service provider had weight 1/2 while all others had weight 1. If the census tract had two service providers each with 7% or less of the total residential subscribers in the census tract, those two service providers each had weight 1/3 while all others had weight 1.[15]

Survey Data Collection

The Bureau contacted each provider that had been selected in the sampling stage. Each provider was asked to report rates for all standalone broadband plans in one or more census tracts. These providers were asked to report these rates via a specially-designed online system for which each provider was given login access. If a provider did not currently offer residential service in the census tract, the provider would indicate this and otherwise report nothing. Providers reported rates beginning December 17, 2013, continuing for several weeks thereafter.

---

[11] Sampling with replacement means that each draw is made from the entire population of urban census tracts. This means it is possible for the same census tract to be drawn more than once.

[12] Specifically, we defined as "urban" for purposes of the survey all 2010 census tract urban areas and urban clusters that sit within Metropolitan Statistical Areas.

[13] Census tract 12083000802 was selected twice. Census tract 05035030201 was also selected twice.

[14] A plot of residential subscribers versus urban census tracts for broadband service providers nationally gave an indication of two bands above and below 100 residential subscribers per census tract. The national average number of broadband residential subscribers in an urban census tract is about 1400; so, 100 residential subscribers represents about 7% of residential subscribers in an average urban census tract.

[15] We had no data that would enable us to weight carriers by their coverage of the tract. Consequently, this weighting scheme is intended to approximate the expected coverage of each sampled carrier in its corresponding residential tract. The approach assumes that if the carrier claims more than 7% of the tract's subscribers, then it covers the tract, but if it uniquely claims less than 7% then it only covers half the tract, and if two carriers each claim less than 7%, then each only covers one third of the tract.

Analysis of the Collected Data

Data Preparation

The Bureau received responses for 498 census tracts from 81 service providers.[16] A total of 2211 rows of data were recorded. A total of 63 rows did not provide monthly rate data, for the following reasons:

- The row gave no indication that the census tract was served by the provider (54)
- The row was an erroneous entry (4)
- The row indicated service at a specified level was provided but no rates were given (3)
- The row indicated that service would be provided at a higher level in the future (1)
- The row was a duplicate entry (1)

In two separate cases identical rates were provided for the same service for the same provider in the same census tract;[17] in each of these two cases, the two duplicate rows were merged into a single row. In addition, some service providers offered the same service in a census tract using digital subscriber line (DSL) and fiber to the home (FTTH) technologies reporting rates for each technology on separate rows. There were 41 such cases where the two rows were merged by averaging the rates for DSL and FTTH technologies.[18] As a result, a total of 2105 monthly rates for broadband service were provided by 71 providers for 444 census tracts.

Values for reported download speeds ranged from 0.5 to 20480 and values for reported upload speeds ranged from 0.125 to 1024. All values were expected to be entered in Mbps, but some respondents evidently entered the relevant data as Kbps. For consistency, speed values entered in the survey were converted as shown in the table below:

Table 1. Speed Conversions

Speed Entered			Speed
0.256	or	256	0.25
0.384	or	384	0.375
0.512	or	512	0.5
0.768	or	768	0.75
1.024	or	1024	1
20.48	or	20480	20

[16] IATD did attempt to follow up with non-responders multiple times. However, IATD was unable to reach anyone who could answer our questions. Given the small number of missing observations, we do not expect inclusion of the data (if available) would have a large impact on the results.

[17] For the purposes of this analysis, a service offering is defined by download speed, upload speed, and usage allowance.

[18] If the same service was provided via FTTH and DSL technologies, the rates were averaged to obtain a single rate for the service offered by the provider. In most cases, the rates were the same (so in effect the two entries were merged into a single entry).

The rates presented below represent the sum of the Monthly Charge, Surcharge, and Other Mandatory Charge (if any) reported by the respondents. In cases where a maximum and minimum charge was provided by the respondent, the average of the maximum and minimum was used.

Two service offering rates from Nitelog Inc were excluded from the analysis as apparent outliers.[19] The rates were $1,250 and $1,999 for 25/25/Unlimited and 50/50/Unlimited using Fixed Wireless technology. The next highest reported monthly rate was $399.95 for 505/100/250 service.

One service offering from Digis LLC for 5/5/Unlimited service using Fixed Wireless technology at a monthly rate of $271.45 was also excluded from the analysis as an apparent outlier. The next highest reported monthly rate for 5/x/Unlimited service was $87.45 for 0.75 Mbps upload speed. The third highest reported monthly rate for 5/x/Unlimited service was $61.45 for 2 Mbps upload speed which was also offered by Digis LLC.

Potential Options

The goal is to develop an approach for determining an upper range of rates that could be reasonably comparable to the national average urban rate for similar broadband services. For purposes of the following discussion, the Bureau defined "similar services" as those with a download speed, upload speed, and usage allowance close to the minimum performance specifications of a download speed of 4 Mbps, an upload speed 1 Mbps, and a usage allowance of 100 GB per month. We note, however, that the options presented could be adapted for use with services offering differing speeds and/or usage allowances and thus would be flexible enough to take account any changes the Commission may make in the future regarding broadband performance obligations.[20]

The following analysis explicitly does *not* select a specific methodology or benchmark. Rather, we present several potential methodologies for determining an upper range that could be adopted by the Bureau at a future date as a benchmark and discuss the benefits and challenges of each. The selection of a method and a value to select with that method are decisions that will be made after further public comment.

The first method is to calculate relatively simple rate statistics for specified subsamples; for example, all rates for observations with the specified download speed, or all rates for observations from providers that offer a service that meets or exceeds a minimum service level. Both of these approaches have the disadvantage of including and/or excluding observations that are close, but not identical to the specified broadband service requirement. A variant of these approaches would be to develop an average rate for a selection of similar services, while testing how sensitive the resulting range is to any given choice of

[19] As part of the determination that these observations should be treated as outliers, we visited the provider's website and determined the service was essentially a business service and not a good comparison for residential service.

[20] The Commission recently proposed in a Further Notice of Proposed Rulemaking to adopt a minimum service standard of 10 Mbps downstream, 1 Mbps upstream, and a 100 GB monthly usage allowance that would apply to all eligible telecommunications carriers that are subject to broadband public interest obligations. The Commission has not, to date, specified a minimum usage allowance for rate-of-return carriers.

similar services. A third approach uses regression analysis to account for the multiple dimensions of broadband service (i.e. download bandwidth, upload bandwidth, and usage allowance).[21]

As a general note, in each methodology, we only present in the main body of the text the point estimates.[22] However, it is important to remember that each point estimate has a statistical error and therefore has a confidence interval around it. Thus, if the statistical error is known, we could say with 95% confidence that the population value lay within a specific interval of its estimate from the sample.

Rate Estimates for Services with the Specified Download Speed

The first approach we consider is the estimation of candidate benchmark values directly from rates from those observations for the specified download speed. Under this approach, we would specify the relevant download speed, say, 4 Mbps, and the relevant cutoff, say, the sample average plus two standard deviations. If rates were normally distributed,[23] this upper bound would represent an unbiased estimate of the rate that was higher than 97.5% of all rates with the download speed of interest. For the reasons discussed below, we would not recommend this approach. However, it has expositional value because it illustrates both the nature of our sample and the problems in trying to define an upper range of rates.

Table 2 below provides estimates of monthly broadband rate statistics for different download speeds or download speed groups.[24] "Responses" is the number of responses out of the 498 received used in the estimate. "Number of Providers" is the number of different providers represented in the observations. All of the remaining seven columns starting with "Median Rate ($)" contain weighted estimates;[25] for each observation, the weight used was the sum of the weights described earlier for service providers in the census tract of the observation.[26] These weights were used in all methodologies described in this document. "% with Unlimited Usage Allowance" is the weighted estimated percentage of offers for services at the specified speed that have an unlimited usage allowance.[27] In Table 2 we present statistics combining all observations for services with download bandwidths between 3 and 4 Mbps. For the combined 3 through 4 Mbps grouping, the mean plus two standard deviations value is $73.22.

[21] Other methods that could be used include quantile regression analysis, non-parametric kernel estimation, and a parameter regression search.

[22] For example, given a sample of three prices ($75, $100, and $125), the average ($100) is a point estimate of the mean.

[23] The rate data has properties (e.g., it is skewed and lumpy) which are not characteristic of a normal distribution.

[24] In Table 2, the download speed groups include fractional speeds. For example, rates for download speeds of 4.5 Mbps are included in the 3 – 4 Mbps speed group and rates for 7.1 Mbps are included in the 7 Mbps speed group.

[25] The weights used are described above on page 3 (for further detail, see the appendix).

[26] Two observations in the sample each represented two sampling units. In these cases, the observation weight was doubled.

[27] In the case of the usage allowance, each observation is a 0 (with a finite usage allowance) or 1 (unlimited usage) and the reported value is the weighted average of the 1's and 0's.

Table 2. Rate Estimates within Download Speed Bands

Download Speed (Mbps)	Number of Providers	Responses	Median Rate($)	Average Rate($)	Std Dev Rate($)	Ave+2SD Rate($)	95% Quantile ($)	97.5% Quantile ($)	% with Unlimited Usage Allowance
0 - 2	28	236	39.78	40.59	10.92	62.43	53.99	69.99	38%
3 - 4	45	242	44.99	47.48	12.87	73.22	64.99	64.99	50%
5	12	67	45.99	46.32	7.27	60.85	59.95	61.45	23%
6	14	125	49.95	48.78	7.60	63.98	50.94	58.97	23%
7	5	33	45.99	48.37	4.94	58.24	54.95	69.49	20%
8	4	17	50.94	57.38	19.27	95.93	95.00	95.00	29%
9	2	2	62.99	63.82	1.44	66.71	66.32	66.32	100%
10	18	47	52.00	58.84	17.44	93.72	99.00	121.45	76%
11 - 15	34	154	55.99	60.56	15.67	91.90	74.99	74.99	78%
16 - 25	26	309	64.95	61.19	14.95	91.10	75.94	96.00	29%
26 - 50	43	292	76.95	86.03	21.17	128.37	115.99	149.00	54%
51 - 100	27	104	94.99	102.45	33.63	169.70	123.00	200.29	87%
101 - 150	18	162	114.95	123.76	16.79	157.34	144.99	144.99	40%
151 - 1000	13	75	304.99	281.91	69.52	420.95	399.95	399.95	82%

The key drawback of this approach is that it only takes into consideration one dimension of the service (i.e. download bandwidth) even though *a priori* we would expect upload bandwidth and usage allowances also to be reflected in the price (for example, this approach would average together a 4/0.4/10 service with a 4/4/1,000 service, if both of those existed). The benefit of this approach, if not its practical usefulness, is that it is straightforward and easily understandable.

Rate Estimates for Service Offerings Meeting or Exceeding a Minimum Service Level

Another approach that focuses on urban rates that meet or exceed a specified minimum service level (MSL) would be to compute the average of the minimal monthly rate for each service provider that meets or exceeds the MSL. To illustrate this approach, a subset of the sample was created consisting of all rates for offerings that met or exceeded the MSL. Then, from this subset, the lowest monthly rate was found for each service provider. For each provider, each census tract with service offered at the provider's lowest rate was included in the estimate. The following table presents estimates of several statistics for monthly service rates based on the observations selected as described above with MSL=4/1/100 and for MSL=10/1/100.

Table 3. Rate Estimates for Service Offerings Meeting or Exceeding a Minimum Service Level

MSL	Providers	Observations	Median	Average	Ave+2SD	97.5% Quantile
4/1/100	64	353	$ 49.95	$ 54.54	$ 82.00	$ 89.00
10/1/100	59	255	$ 54.99	$ 58.05	$ 84.15	$ 79.95

The benefit of this approach is its simplicity and that it includes all providers offering service meeting or exceeding the MSL. The negatives of this approach are that:

- it incorporates observations into the benchmark for urban services with characteristics that are far above the MSL, which are not "similar" services; and
- it may exclude services that are very close to, but do not quite meet the MSL.

A More General Approach to Selecting Sub-samples

Both of the approaches just examined involve the selection of sub-samples for analysis (all those rates for services that deliver the minimum download speed, and the minimum rate for each provider that has at least one service that meets or exceeds the MSL). However, in both cases observations below the MSL (or its proxy) are excluded. A variation on these approaches is to include observations for offerings with differing characteristics within a certain range or ranges below the chosen MSL as well as above the MSL. The challenge of doing so, however, is deciding what is the appropriate range that should be deemed "similar" to the specified performance standard.

Rate Estimates from a Weighted Linear Regression Model

The third approach is based on a weighted[28] linear regression model.[29] This has an important advantage over the use of simple averages in that it provides a formalized means of estimating the various degrees to which the different service characteristics (download speed, upload speed, and usage allowance) influence rates. However, it also requires similar decisions to those made above. Because inclusion of observations from services dramatically different from a MSL plan might influence the ultimate benchmark, it may be appropriate to use a subsample, that is, to fit a model using data only in the region of interest for the MSL. In particular, we found that standard deviations of rates with less than 15 Mbps download speed tend to be smaller than those at higher download speeds. Consequently, using a model fitting all the data as opposed to one fitting data using observations in the lower range of speeds could result in overestimation of the standard deviation appropriate to the MSL and consequently also the benchmark rate.

[28] Recent research has investigated issues with weighting (see for example Gary Solon, Steven J. Haider, and Jeffrey Wooldridge, "What are We Weighting For?" NBER Working Paper No. 18859, February 2013). The use of weights here is consistent with the estimation described in the Appendix for adjustments based on the sampling procedure. An unweighted regression was performed. For the 4/1/100 level of service, the mean estimate was $0.54 lower than when using the weights.

[29] We did investigate other regression specifications. While some of these showed small improvements in the goodness of fit, the improvement did not justify the increased complexity of the model. While goodness of fit is one element of model evaluation, it is also important that the model make sense for the actual application, in this case how three service characteristics may influence broadband rates. Since, in this case, there is no obvious reason for more complex price determination models, we narrowed our focus to the simple linear case.

To illustrate this approach, we applied a multidimensional weighted linear regression technique to all services with download bandwidths of 15 Mbps or less.[30] This sub-sample of the data encompassed 995 rates from 65 different providers. The rates in this sub-sample ranged from $11.46 to $151.45 with a weighted standard deviation of $14.22. We undertook a weighted linear regression fit based on the following model:

$$\text{Average Monthly Rate (\$)} = K_0 + K_D\, D + K_U\, U - K_A\, A$$

for download speed in Mbps (D), upload speed in Mbps (U), and usage allowance in GB (A = 1/UsageAllowance or 0 if unlimited usage)[31] was used.[32] We estimated the parameters as:

$$\text{Average Monthly Rate (\$)} = 41.247 + 1.02463\, D + 2.75597\, U - 335.676\, A.$$

The weighted R Squared was 0.30 and each estimated coefficient was significant at the 0.1% confidence level.

The table below shows the model's average monthly rate estimates for various service levels.

Table 4. Estimates of Average Monthly Rate Based on the Linear Regression Model

Speed (Mbps)	Usage Allowance (GB)		
Down/Up	100	250	No Limit
3/.5	$ 42.34	$ 44.36	$ 45.70
3/1	$ 43.72	$ 45.73	$ 47.08
4/1	$ 44.74	$ 46.76	$ 48.10
5/.5	$ 44.39	$ 46.41	$ 47.75
5/1	$ 45.77	$ 47.78	$ 49.13
6/.5	$ 45.42	$ 47.43	$ 48.77
6/1	$ 46.79	$ 48.81	$ 50.15
10/1	$ 50.89	$ 52.91	$ 54.25

[30] The selection of the subset of data for use in the regression analysis was guided by the current 4 and proposed 10 Mbps minimum download speed service requirement. The subset had to be large enough to encompass these candidates. However, since Table 2 indicates a trend of increasing variability in rates with increasing download speed, an estimate of the standard deviation in the neighborhood of these candidates would be more accurate if the subset did not extend far beyond 10 Mbps. A subset of 0 -15 Mbps placed the candidates in the center of the data range where regression estimates are more accurate. A different subset would likely be called for if a different range of candidates were envisioned.

[31] The inverse of the usage allowance was used in the model to accommodate services with unlimited usage. For such services, the usage allowance is infinity, which cannot be used in the regression model. When the inverse of the usage allowance is used, unlimited usage is represented by 0 which can be used in the regression model. A regression fit can be made using usage allowances directly if an additional dummy variable is included indicating whether usage is limited, the results of which do not materially differ from the regression using the inverse (see footnote 34 below).

[32] Because we were focused only on the service characteristics, we did not include a variable to capture the technology over which the service is provided.

The table below shows the standard deviation of error for the average monthly rate estimates in Table 4.

Table 5. Standard Deviation of Error in Estimates of Average Monthly Rate in Table 4

Speed (Mbps)	Usage Allowance (GB)		
Down/Up	100	250	No Limit
3/.5	$ 0.71	$ 0.44	$ 0.57
3/1	$ 0.74	$ 0.45	$ 0.57
4/1	$ 0.73	$ 0.40	$ 0.52
5/.5	$ 0.74	$ 0.43	$ 0.54
5/1	$ 0.73	$ 0.39	$ 0.49
6/.5	$ 0.78	$ 0.47	$ 0.56
6/1	$ 0.75	$ 0.40	$ 0.48
10/1	$ 0.96	$ 0.65	$ 0.65

A 95% confidence interval for the estimates in Table 4 would be roughly +/- twice the values in Table 5.

Various quantile levels can be estimated using the following table with the equation

$$\text{Monthly Rate Quantile P} = \text{Average Monthly Rate} + Q_P \text{ SD}$$

where SD is the weighted standard deviation about the regression fit ($11.87).[33]

Table 6. Quantiles of the Standard Normal Distribution

P	Q_P
90%	1.282
95%	1.645
97.5%	1.960
99%	2.326

Using the equation above, the table below shows the model's average monthly rates plus twice the standard deviation for the same set of service levels as in Table 4; these values are roughly the 97.5% quantiles for the rates.

[33] Fitting all of the data using the model would have resulted in a larger value for SD which would have been less representative of the spread in the rate values in the vicinity of the candidate MSLs.

Table 7. Estimates of Average Monthly Rate Plus 2 Standard Deviations
Based on the Linear Regression Model

Speed (Mbps)	Usage Allowance (GB)		
Down/Up	100	250	No Limit
3/.5	$ 66.08	$ 68.10	$ 69.44
3/1	$ 67.46	$ 69.47	$ 70.82
4/1	$ 68.48	$ 70.50	$ 71.84
5/.5	$ 68.13	$ 70.15	$ 71.49
5/1	$ 69.51	$ 71.52	$ 72.87
6/.5	$ 69.16	$ 71.17	$ 72.51
6/1	$ 70.53	$ 72.55	$ 73.89
10/1	$ 74.63	$ 76.65	$ 77.99

For example, using the above estimated regression model to set a broadband reasonable comparability benchmark for the minimum service characteristics based on the average rate plus twice the standard deviation:

- If the minimum broadband performance standard is 4/1 Mbps with a 100 GB usage allowance, then the reasonable comparability benchmark would be $68.48.
- If the minimum broadband performance standard is 10/1 Mbps with a 100 GB usage allowance, then the reasonable comparability benchmark would be $74.63.[34]

Not surprisingly, these numbers are lower than the results of the second approach which includes observations that exceed the specified minimum service standard. These estimates from linear regression take into account various service characteristics, while the previous approach utilized observations for services with differing service characteristics without adjusting for those characteristics. We note, however, these are only examples.

[34] If the regression model had used usage allowance (adding a dummy variable indicating whether usage is limited), rather than its inverse, the benchmark rate for 4/1/100 service would be $68.73, and for 10/1/100, $75.00.

Appendix: Technical Background

The sample process was designed to estimate the mean and standard deviation of the distribution of available service rates for broadband service in urban areas. These estimates could then be used as input for establishing benchmarks; for example, the mean plus twice the standard deviation is a possible upper limit based on the approximate 97.5 percentile of a normal distribution.

At a conceptual level, the "distribution of available service rates in urban areas" could be captured through the following process:

1. For each household in an urban area in the United States, list all the service providers offering fixed broadband service to that household and the service rates they offer for each level of service.

2. Concatenate all the lists from each household into a single list.

The resulting list of rates is the distribution of available service rates in urban areas for fixed broadband service at various levels of service.

If we were to focus on the rates for a specific level of service, the mean and standard deviation of available rates would be

$$\bar{R} = \sum_{i=1}^{N}\sum_{j=1}^{J_i} R_{ij}/N_R$$

$$\sigma_R = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left(R_{ij}-\bar{R}\right)^2/N_R}$$

where

R_{ij} = jth rate available to household i

J_i = number of rates available to household i

N = number of eligible households

N_R = Total number of available rates = $\sum_{i=1}^{N} J_i$

From a practical standpoint, an equivalent result may be obtained by surveying service providers offering the relevant service in urban areas to obtain data on their rates. In this frame, the equivalent mean of the distribution of available rates is obtained as the weighted sum of rates offered by service providers in each census tract. Similarly, the equivalent standard deviation of the distribution of available rates is obtained as the square root of the weighted sum of squared differences between the mean rate of the distribution and rates offered by service providers in each census tract.

$$\bar{R} = \sum_{i=1}^{U} \sum_{k=1}^{K_i} W_{ik} Y_{ik} / N_R$$

$$\sigma_R = \sqrt{\sum_{i=1}^{N} \sum_{k=1}^{K_i} W_{ik} (Y_{ik} - \bar{R})^2 / N_R}$$

where

Y_{ik} = rate offered in census tract i by service provider k

W_{ik} = number of households in census tract i offered service by service provider k

K_i = number of service providers offering service in census tract i

U = number of urban census tracts

W_i = Total number of available rates in census tract i = $\sum_{k=1}^{K_i} W_{ik}$

N_R = Total number of available rates = $\sum_{i=1}^{U} W_i$

In order to estimate the mean and the standard deviation, a sample of service providers offering fixed broadband service were surveyed for rates they offer in a sample of urban census tracts. The sampling process was as follows:

- A census tract i was randomly selected with probability H_i/H where H_i is the number of households in census tract i and H is the sum of the H_i over all census tracts.
- A carrier k was randomly selected from the K_i carriers offering service in census tract i with probability W_{ik} / W_i
- This process is repeated n=500 times to obtain 500 sampling units. We note that sampling units could appear multiple times in the sample.

The mean of the rate distribution was estimated as the ratio of total dollars in rate offers to the total number of rates. We note that the total number of available rates is not known, so it must be estimated from the sample as well as the estimate of total dollars in rate offers. Consequently, an estimate of the mean of available rates based on this sample is

$$\ddot{\bar{R}} = \left(\frac{\sum_{j=1}^{n} x_j / P_j}{n}\right) / \left(\frac{\sum_{j=1}^{n} z_j / Q_j}{n}\right)$$

where

$X_j = W_{ik}\, Y_{ik}$ from the j[th] sampling unit (census tract i and carrier k),

P_j = probability of selecting the jth sampling unit = $(H_i/H)(W_{ik} / W_i)$ for the j[th] sampling unit,

$Z_j = W_i$ from the j[th] sampling unit,

Q_j = probability of selecting the jth urban area = H_i/H

The estimate of the mean can be simplified to

$$\ddot{\bar{R}} = \frac{\sum_{j=1}^{n} F_j Y_j}{\sum_{j=1}^{n} F_j}$$

where Y_j is the rate Y_{ik} and F_j is W_i/H_i for the jth sampling unit.

The values for the W_i are not known. As described in the main text, weights between 0 and 1 were assigned to carriers in each census tract of the sample based on their share of residential subscribers in the tract. These weights are expressions of W_{ik}/H_i (the fraction of households carrier k offers service in census tract i) and therefore F_i is the sum of these weights for carriers in census tract i. Similarly, the estimate of the standard deviation is

$$\widehat{\sigma_{\bar{R}}} = \sqrt{\frac{\sum_{j=1}^{n} F_j \left(x_j - \hat{\bar{R}}\right)^2}{\sum_{j=1}^{n} F_j}}.$$